EXHIBIT 4

Ballot-marking devices (BMDs) cannot assure the will of the voters

Andrew W. Appel[†]
Princeton University

Richard A. DeMillo[†]
Georgia Tech

Philip B. Stark[†]
Univ. of California, Berkeley

April 21, 2019

Abstract

Computers, including all modern voting systems, can be hacked and misprogrammed. The scale and complexity of U.S. elections may require the use of computers to count ballots, but election integrity requires a paper-ballot voting system in which, regardless of how they are initially counted, ballots can be recounted by hand to check whether election outcomes have been altered by buggy or hacked software. Furthermore, secure voting systems must be able to recover from any errors that might have occurred.

However, paper ballots provide no assurance unless they accurately record the vote as the voter *expresses* it. Voters can express their intent by hand-marking a ballot with a pen, or using a computer called a ballot-marking device (BMD), which generally has a touchscreen and assistive interfaces. Voters can make mistakes in *expressing* their intent in either technology, but only the BMD is *also* subject to systematic error from computer hacking or bugs in the process of recording the vote on paper, after the voter has expressed it. A hacked BMD can print a vote on the paper ballot that differs from what the voter expressed, or can omit a vote that the voter expressed.

It is not easy to check whether BMD output accurately reflects how one voted in every contest. Research shows that most voters do not review paper ballots

[†]Authors are listed alphabetically; they contributed equally to this work.

printed by BMDs, even when clearly instructed to check for errors. Furthermore, most voters who do review their ballots do not check carefully enough to notice errors that would change how their votes were counted. Finally, voters who detect BMD errors before casting their ballots, can correct only their own ballots, not systematic errors, bugs, or hacking. There is no action that a voter can take to demonstrate to election officials that a BMD altered their expressed votes, and thus no way voters can help deter, detect, contain, and correct computer hacking in elections. That is, not only is it inappropriate to rely on voters to check whether BMDs alter expressed votes, *it doesn't work*.

Risk-limiting audits of a trustworthy paper trail can check whether errors in tabulating the votes *as recorded* altered election outcomes, but there is no way to check whether errors in how BMDs record *expressed* votes altered election outcomes. The outcomes of elections conducted on current BMDs therefore cannot be confirmed by audits. This paper identifies two properties of voting systems, *contestability* and *defensibility*, that are necessary conditions for any audit to confirm election outcomes. No commercially available EAC-certified BMD is contestable or defensible.

To reduce the risk that computers undetectably alter election results by printing erroneous votes on the official paper audit trail, the use of BMDs should be limited to voters who require assistive technology to vote independently.

Elections for public office and on public questions in the United States or any democracy must produce outcomes based on the votes that voters *express* when they indicate their choices on a paper ballot or on a machine. Computers have become indispensable to conducting elections, but computers are vulnerable. They can be hacked—compromised by insiders or external adversaries who can replace their software with fraudulent software that deliberately miscounts votes—and they can contain design errors and bugs—hardware or software flaws or configuration errors that result in mis-recording or mis-tabulating votes. Therefore there must be some way, *independent* of any software in any computers, to ensure that reported election outcomes are correct, i.e., consistent with the expressed votes as intended by the voters.

Voting systems should be *software independent*, meaning that "an undetected change or error in its software cannot cause an undetectable change or error in an election outcome" [23]. Indeed, version 2.0 of the Voluntary Voting System Guidelines (VVSG 2.0) incorporates this principle [7].

Software independence is similar to tamper-evident packaging: if somebody opens the container and disturbs the contents, it will leave a trace.

While software independence is crucial, it is not enough: *who* can detect errors and *what happens* when errors are detected are just as important. Even if individual voters in principle could detect changes to their votes on the BMD-generated ballot, unless voters can provide convincing evidence of problems to the public and unless election officials take appropriate remedies when presented with such evidence, software independence alone does not guarantee that outcome-changing problems—accidental or malicious—can be caught, much less corrected.

To be acceptable, a voting system also must be *contestable*: We say that voting system is contestable if any change or error in its software that results in a change or error in a reported election outcome can generate public evidence that the reported outcome is not trustworthy. Evidence available only to individual voters[1] does not suffice: "trust me" is not evidence. If a voting system is contestable, it is software independent, but the converse is not necessarily true. If a voting system is not contestable, then problems might never see the light of day, much less be corrected.

Voting systems must also be *defensible*. We say that a voting system is defensible if, when it reports the correct outcome, it can also generate convincing public evidence that the reported outcome is correct. Evidence available only to an election official or voting system vendor does not suffice: in other words, "trust me" is not evidence. If a voting system is not defensible, then it is vulnerable to "crying wolf": malicious actors could claim that the system malfunctioned when in fact it did not, and election officials will have no way to prove otherwise.

Rivest and Wack [23] also define a voting system to be *strongly software independent* if it is software independent and moreover, a detected change or error in an election outcome (due to change or error in the software) can be corrected using only the ballots and ballot records of the current election.[2] Strong software independence combines tamper evidence with a kind of resilience: there's a way to tell whether faulty software caused a problem, and a way to recover from the problem if it did.

The only known practical technology for a contestable, defensible, strongly software independent voting system is *hand-marked paper ballots*, kept physically secure,

[1] Specifically, if the voter is selected candidate A on the touchscreen of a BMD, but the BMD prints candidate B on the paper ballot, then this A-vs-B evidence is available to the individual voter, but the voter cannot demonstrate this evidence to anyone else, since nobody else saw—nor should have seen—where the voter touched the screen. Thus, the voting system cannot generate public evidence of errors recording expressed votes, even if those errors altered the reported outcome.

[2] The only alternative remedy would be to void the results of the entire election and conduct a new one.

counted by machine, audited manually, and recountable by hand.[3]

Over 40 states now use some form of paper ballot for most voters [14]. Most of the remaining states are taking steps to adopt paper ballots. But *not all voting systems that use paper ballots are equally secure*. Some are not software independent. Some are software independent but not contestable or defensible. In this report we explain:

- *Hand-marked paper ballot* systems are the only practical technology for contestable, defensible voting systems.
- *Some ballot-marking devices (BMDs)* can be software independent, but they are neither contestable nor defensible. Hacked or misprogrammed BMDs can alter election outcomes undetectably, and elections conducted using BMDs do not provide public evidence that reported outcomes are correct. Therefore BMDs should not be used by voters who are able to mark an optical-scan ballot with a pen.
- *All-in-one BMD or DRE+VVPAT voting machines* are not software independent, contestable, or defensible. They should not be used in public elections.

Terminology

Although a voter may form an intention to vote for a candidate or issue days, minutes, or seconds before actually casting a ballot, that intention is a psychological state that cannot be directly observed by anyone else. Others can have access to that intention through what the voter (privately) *expresses* to the voting technology by interacting with it, e.g., by making selections on a BMD or marking a ballot by hand.[4] Voting systems must accurately record the vote as the voter *expressed* it.

With a *hand-marked paper ballot optical-scan* system, the voter is given a paper ballot on which all choices (candidates) in each contest are listed; next to each candidate

[3]The election must also generate convincing evidence that physical security of the ballots was not compromised, and the audit must generate convincing public evidence that the audit itself was conducted correctly.

[4]We recognize that voters make mistakes in expressing their intentions. For example, they may misunderstand the layout of a ballot or through a perceptual error or lapse of attention make an unintended choice. The use of touchscreen technology does not necessarily correct for such user errors, as every smartphone user who has mistyped an important text message knows. Poorly designed ballots, poorly designed touchscreen interfaces, and poorly designed assistive interfaces increase the rate of error in voters' expressions of their votes. For the purposes of this report, we assume that properly engineered systems seek to minimize such usability errors.

is a *target* (typically an oval or other shape) which the voter marks with a pen to indicate a vote. Ballots may be either preprinted or printed (unvoted) at the polling place using *ballot on demand* printers. In either case, the voter creates a tamper-evident record of intent by marking the printed paper ballot with a pen.

Such hand-marked paper ballots may be scanned and tabulated at the polling place using a *precinct-count optical scanner* (PCOS), or may be brought to a central place to be scanned and tabulated by a *central-count optical scanner* (CCOS). Mail-in ballots are typically counted by CCOS machines.

After scanning a ballot, a PCOS machine deposits the ballot in a secure, sealed ballot box for later use in recounts or audits; this is *ballot retention*. Ballots counted by CCOS are also retained for recounts or audits.[5]

Paper ballots can also be hand counted, but in most jurisdictions (especially where there are many contests on the ballot) this is hard to do quickly; Americans expect election-night reporting of unofficial totals. Hand counting—i.e., manually determining votes directly from the paper ballots—is appropriate for audits and recounts.

A *ballot-marking device* (BMD) provides a computerized user interface that presents the ballot to voters and captures their expressed selections, for instance, a touchscreen interface or an assistive interface that enables voters with disabilities to vote independently. Voter inputs (expressed votes) are recorded electronically. When a voter indicates that the ballot is complete and ready to be cast, the BMD prints a paper version of the electronically marked ballot. We generally use the term *BMD* for devices that mark ballots but do not tabulate or retain them, and *all-in-one* for devices that combine ballot marking, tabulation, and retention into the same paper path.

The paper ballot printed by a BMD may be in the same format as an optical-scan form (e.g., with ovals filled as if by hand) or it may list just the names of the candidate(s) selected in each contest. The BMD may also encode these selections into barcodes or QR codes for optical scanning. We discuss issues with barcodes later in this report.

An *all-in-one touchscreen voting machine* combines computerized ballot marking, tabulation, and retention in the same paper path. All-in-one machines come in several configurations:

- DRE+VVPAT machines—direct-recording electronic (DRE) voting machines with a voter-verifiable paper audit trail (VVPAT)—provide the voter a touchscreen (or

[5]Regulations and procedures governing custody and physical security of ballots are uneven and in many cases inadequate, but simple to correct because of decades of development of best practices.

other) interface, then print a paper ballot that is displayed to the voter under glass. The voter is expected to review this ballot and approve it, after which the machine deposits it into a ballot box. DRE+VVPAT machines do not contain optical scanners; that is, they do not read what is marked on the paper ballot; instead, they tabulate the vote directly from inputs to the touchscreen or other interface.

- BMD+Scanner all-in-one machines[6] provide the voter a touchscreen (or other) interface to input ballot choices and print a paper ballot that is ejected from a slot for the voter to inspect. The voter then reinserts the ballot into the slot, after which the all-in-one BMD+scanner scans it and deposits it into a ballot box.

Opscan+BMD with separate paper paths. At least one model of voting machine (the Dominion ICP320) contains an optical scanner and a BMD in the same cabinet,[7] so that the optical scanner and BMD-printer are not in the same paper path; no possible configuration of the software could cause a BMD-marked ballot to be deposited in the ballot box without human handling of the ballot. We do not classify this as an *all-in-one* machine.

Hacking

There are many forms of computer hacking. In this analysis of voting machines we focus on the alteration of voting machine software so that it miscounts votes or mismarks ballots to alter election outcomes. There are many ways to alter the software of a voting machine: a person with physical access to the computer can open it and directly access the memory; one can plug in a special USB thumbdrive that exploits bugs and vulnerabilities in the computer's USB drivers; one can connect to its WiFi port or Bluetooth port or telephone modem (if any) and exploit bugs in those drivers, or in the operating system.

"Air-gapping" a system (which is to say, disconnecting it from a wired network) does not automatically protect it. Before each election, election administrators must transfer a *ballot definition* into the voting machine by inserting a *ballot definition cartridge* that was programmed on election-administration computers that may have been connected previously to various networks; it has been demonstrated that vote-changing viruses can propagate via these ballot-definition cartridges [13].

Hackers might be corrupt insiders with access to a voting-machine warehouse; cor-

[6] The ES&S ExpressVote can be configured as either a BMD or a BMD+Scanner all-in-one.

[7] More precisely, the ICP320 optical scanner and the BMD audio+buttons interface are in the same cabinet, but the printer is a separate box.

rupt insiders with access to a county's election-administration computers; outsiders who can gain remote access to election-administration computers; outsiders who can gain remote access to voting-machine manufacturers' computers (and "hack" the firmware installed in new machines, or the firmware updates supplied for existing machines), and so on. Supply-chain hacks are also possible: the hardware installed by a voting system vendor may have malware pre-installed by the vendor's component suppliers.[8]

Computer systems (including voting machines) have so many layers of software that it is impossible to make them perfectly secure [18, pp. 89–91]. When manufacturers of voting machines use the best known security practices, adversaries may find it more difficult to hack a BMD or optical scanner—but not impossible. Every computer in every critical system is vulnerable to compromise through hacking, insider attacks or exploiting design flaws.

Election assurance through risk-limiting audits.

To ensure that the reported outcome of each contest is that outcome that would have been found by accurately tabulating the voters' intent as recorded, the most practical known technology is a *risk-limiting audit* (RLA) of paper ballots [25, 26, 17]. The National Academies of Science, Engineering, and Medicine, recommend routine RLAs after every election [18], as do many other organizations and entities concerned with election integrity.[9]

A RLA involves manually inspecting randomly selected paper ballots following a rigorous protocol. The audit stops if and when the sample provides convincing evidence that the reported outcome is correct; otherwise, the audit continues until every ballot has been inspected manually and the correct electoral outcome is known.

RLAs can check whether errors in tabulating recorded votes altered election outcomes, but cannot check whether errors in recording expressed votes altered election outcomes. Properly preserved hand-marked paper ballots ensure that expressed votes are identical to recorded votes. On the other hand, BMDs might not record expressed

[8] Given that many chips and other components are manufactured in China and elsewhere, this is a serious concern. Carsten Schürmann has found Chinese pop songs on the internal memory of voting machines (C. Schürmann, personal communication, 2018). Presumably those files were left there accidentally—but this shows that malicious code *could* have been pre-installed deliberately, and that neither the vendor's nor the election official's security and quality control measures discovered and removed the extraneous files.

[9] Among them are the Presidential Commission on Election Administration, the American Statistical Association, the League of Women Voters, and Verified Voting Foundation.

votes accurately, for instance, if BMD software has bugs, was misconfigured, or was hacked. Thus, RLAs that rely on BMD output cannot ensure that election outcomes are correct.

RLAs protect against vote-tabulation errors, whether those errors are caused by failures to follow procedures, misconfiguration, miscalibration, faulty engineering, bugs, or malicious hacking.[10] The *risk limit* of a risk-limiting audit is the maximum chance that an outcome that is incorrect because of tabulation errors will pass the audit without being corrected. The risk limit should be determined as a matter of policy or law. For instance, a 5% risk limit means that, if a reported outcome is wrong because of tabulation errors, there is at least a 95% chance that the post-election audit will correct it. Smaller risk limits give higher confidence in election outcomes, but require inspecting more ballots, other things being equal. RLAs never revise a correct outcome.

RLAs can be very efficient, depending in part on how the voting system is designed. If the computer results are accurate, an efficient RLA with a risk limit of 5% requires examining about (7 divided by the margin) ballots selected randomly from the contest.[11] For instance, if the margin of victory is 10% and the results are correct, the RLA would need to examine about $7/10\% = 70$ ballots to confirm the outcome at 5% risk. For a 1% margin, the RLA would need to examine about $7/1\% = 700$ ballots. The sample size does not depend (much) on the total number of ballot cast in the contest, only on the margin of the winning candidate's victory.

A paper-based voting system (such as one that uses optical scanners) is systematically more secure than a paperless system (such as DREs) only if the paper trail is trustworthy and the results are audited against the paper trail using a rigorous method such as an RLA.

But what if the paper ballots are not a trustworthy record of the votes expressed by the voters? If it is possible that error, hacking, bugs, or miscalibration caused the recorded votes to differ from the expressed votes, an RLA or even a full hand recount does not provide convincing public evidence that election outcomes are correct: such a system cannot be *defensible*. In short, paper ballots provide little assurance against hacking if they are never examined or if the paper might not accurately record the vote expressed by the voter.

[10] RLAs do not protect against problems that cause BMDs to print something other than what was shown to the voter on the screen, nor do they protect against problems with ballot custody.

[11] Technically, it is the *diluted margin* that enters the calculation. The diluted margin is the number of votes that separate the winner with the fewest votes from the loser with the most votes, divided by the number of ballots cast, including undervotes and invalid votes.

Security Flaws

A BMD-generated paper trail is not a reliable record of the vote expressed by the voter. Like any computer, a BMD (or a DRE+VVPAT) is vulnerable to hacking, installation of unauthorized (fraudulent) software, and alteration of installed software.[12]

If a hacker sought to steal an election by altering BMD software, what would the hacker program the BMD to do? In cybersecurity practice, we call this the *threat model*.

The simplest threat model is this one: In some contests, not necessarily top-of-the-ticket, change a small percentage of the votes (such as 5%).

In recent national elections, analysts have considered a candidate who received 60% of the vote to have won by a landslide. Many contests are decided by less than a 10% margin. Changing 5% of the votes can change the margin by 10%, because "flipping" a vote for one candidate into a vote for a different candidate changes the difference in their tallies—i.e., the margin—by 2 votes. If hacking or bugs or misconfiguration could change 5% of the votes, that would be a very significant threat.

Although public and media interest often focus on top-of-the-ticket races such as President and Governor, elections for lower offices such as state representatives, who control legislative agendas and redistricting, and county officials, who manage elections and assess taxes, are just as important in our democracy. But most voters are not as familiar with the names of the candidates for those offices.

Research by one of us [9], in a real polling place in Tennessee during the 2018 election, found that half the voters *didn't look at all* at the paper ballot printed by a BMD, even when they were holding it in their hand and directed to do so while carrying it from the BMD to the optical scanner. Those voters who did look at the BMD-printed ballot spent *an average of 4 seconds* examining it to verify that the eighteen or more choices they made were correctly recorded. That amounts to 222 milliseconds per contest, barely enough time for the human eye to move and refocus under perfect conditions and not nearly enough time for perception, comprehension, and recall [22].[13] [14]

[12] It is also vulnerable to bugs and misconfiguration.

[13] You might think, "the voter really *should* carefully review their BMD-printed ballot." But because the scientific evidence shows that voters *do not* [9] and cognitively *cannot* [12] perform this task well, legislators and election administrators should provide a voting system that counts the votes *as voters express them*.

[14] Studies of voter confidence about their ability to verify their ballots are not relevant: in typical situations, subjective confidence and objective accuracy are at best weakly correlated. The relationship between confidence and accuracy has been studied in contexts ranging from eyewitness accuracy [6, 8,

The same study found that among voters who examined their hand-marked ballots, half were unable to recall key features of ballots cast moments before, a prerequisite step for being able to recall their own ballot choices.

Suppose, then, that 10% of voters examine their paper ballots carefully enough to even *see* the candidate's name recorded as their vote for legislator or county commissioner. Of those, perhaps only half will remember the name of the candidate they intended to vote for.[15]

Of those who notice that the vote printed is not the candidate they intended to vote for, what are they supposed to think, and what are they supposed to do? Do they think, "Oh, I must have made a mistake on the touchscreen," or do they think, "Hey, the machine is cheating or malfunctioning!" There's no way for the voter to know for sure—voters do make mistakes—and there's *absolutely* no way for the voter to prove to a pollworker or election official that a BMD printed something other than what the voter entered on the screen.[16]

Either way, polling place procedures generally advise voters to ask a pollworker for a new ballot if theirs does not show what they intended. Pollworkers should void that BMD-printed ballot, and the voter should get another chance to mark a ballot. Anecdotal evidence suggests that many voters are too timid to ask, or don't know that they have the right to ask, or are not sure whom to ask. Even if a voter asks for a new ballot, training for pollworkers is uneven, and we are aware of no formal procedure for resolving disputes if a request for a new ballot is refused. Moreover, there is no sensible protocol for ensuring that BMDs that misbehave are investigated—nor can there be, as we argue below.

Let's summarize. If a machine alters votes on 5% of the ballots (enabling it to change the margin by 10%), then optimistically we might expect $\frac{1}{20} \times \frac{1}{10} \times \frac{1}{2}$ or 0.25% of the voters to request a new ballot and correct their vote. This means that the machine will change the margin by 9.75% and get away with it.

In this scenario, 0.25% of the voters, one in every 400 voters, has requested a new ballot. You might think, "that's a form of *detection* of the hacking." But is isn't, as a

28] to confidence in psychological clinical assessments [10] and social predictions [11]. The disconnect is particularly severe at high confidence. Indeed, this is known as "the overconfidence effect." For a lay discussion, see *Thinking, Fast and Slow* by Nobel economist Daniel Kahnemann [15].

[15] We ask the reader, "do you know the name of the most recent losing candidate for county commissioner?" We recognize that some readers of this document *are* county commissioners, so we ask those readers to imagine the frame of mind of their constituents.

[16] Voters should *certainly* not videorecord themselves voting! That would defeat the privacy of the secret ballot and is illegal in most jurisdictions.

practical matter: a few individual voters may have detected that there was a problem, but there's no procedure by which this translates into any action that election administrators can take to correct the outcome of the election. Polling place procedures *cannot correct or deter hacking, or even reliably detect it*, as we discuss next. This is essentially the distinction between a system that is merely software independent and one that is contestable: a change to the software that alters the outcome might generate evidence for an alert, conscientious, individual voter, but it does not generate public evidence that an election official can rely on to conclude there is a problem.

Even if some voters notice that BMDs are altering votes, there's no way to correct the election outcome. Suppose a state election official wanted to detect whether the BMDs are cheating, and correct election results, based on actions by those few alert voters who notice the error. What procedures could possibly work against the manipulation we are considering?

1. How about, "If at least 1 in 400 voters claims that the machine misrepresented their vote, void the entire election."[17] No responsible authority would implement such a procedure. A few dishonest voters could collaborate to invalidate entire elections simply by falsely claiming that BMDs changed their votes.
2. How about, "If at least 1 in 400 voters claims that the machine misrepresented their vote, then investigate." Investigations are fine, but then what? The only way an investigation can ensure that the outcome accurately reflects what voters expressed to the BMDs is to void an election in which the BMDs have altered votes and conduct a new election. But how do you know whether the BMDs have altered votes, except based the claims of the voters?[18] Furthermore, the investigation itself would suffer from the same problem as above: how can one distinguish between voters who detected BMD hacking or bugs from voters who just want to interfere with an election?

This is the essential security flaw of BMDs: few voters will notice and promptly report discrepancies between what they saw on the screen and what is on the BMD print-

[17]Note that in many jurisdictions, far fewer than 400 voters use a given machine on election day: BMDs are typically expected to serve fewer than 300 voters per day. (The vendor ES&S recommended 27,000 BMDs to serve Georgia's 7 million voters, amounting to 260 voters per BMD [24].) Recall also that the rate 1 in 400 is tied to the amount of manipulation. What if the malware flipped only one vote in 50, instead of 1 vote in 20? That could still change the margin by 4%, but—in this hypothetical—would be noticed by only one voter in 1,000, rather than one in 400. The smaller the margin, the less manipulation it would have taken to alter the electoral outcome.

[18]Forensic examination of the BMD might show that it *was* hacked or misconfigured, but it cannot prove that the BMD *was not* hacked or misconfigured.

out, and even when they do notice, there's nothing appropriate that can be done. (Nor should it be the responsibility of voters to test voting-machine security and accuracy—this is a difficult burden that should not be placed on the voters.)

Therefore, BMDs should not be used by most voters.

Why can't we rely on pre-election and post-election logic and accuracy testing? Most, if not all, jurisdictions perform some kind of *logic and accuracy testing* (LAT) of voting equipment before elections. LAT generally involves voting on the equipment using various combinations of selections, then checking whether the equipment tabulated the votes correctly. As the Volkswagen/Audi "Dieselgate" scandal shows, devices can be programmed to behave properly when they are tested but misbehave in use. Therefore, LAT can never prove that voting machines performed properly in practice.

Don't voters need to check hand-marked ballots, too? It is always a good idea to check one's work. The difference is, with hand-marked paper ballots, voters are responsible for catching and correcting *their own errors*, while if BMDs are used, voters are also responsible for catching *machine errors, bugs, and hacking*. Voters are the *only* people who can detect such problems with BMDs—but, as explained above, if voters do find problems, there's no way they can prove to poll workers or election officials that there were problems and no way to ensure that election officials take appropriate remedial action.

Other tradeoffs, BMDs versus hand-marked opscan

Supporters of ballot-marking devices advance several other arguments for their use.

- **Mark legibility.** A common argument is that a properly functioning BMD will generate clean, error-free, unambiguous marks, while hand marked paper ballots may contain mistakes and stray marks that make it impossible to discern a voter's intent. However appealing this argument seems at first blush, the data are not nearly so compelling. Experience with statewide recounts in Minnesota and elsewhere suggest that truly ambiguous handmade marks are very rare.[19] For instance, 2.9 million hand-marked ballots were cast in the 2008 Minnesota race

[19] States do need clear and complete regulations for interpreting voter marks.

between Al Franken and Norm Coleman for the U.S. Senate. In a manual recount, between 99.95% and 99.99% of ballots were unambiguously marked.[20] [21] In addition, usability studies of hand marked bubble ballots—the kind in most common use in U.S. elections—indicate a *voter* error rate of 0.6%, much lower than the 2.5–3.7% error rate for machine-marked ballots [12].[22] Moreover, modern image-based opscan equipment is better than older "marksense" machines at interpreting imperfect marks. Thus, mark legibility is not a good reason to adopt BMDs for all voters.

- **Undervotes, overvotes.** Another argument offered for BMDs is that the machines can alert voters to undervotes and prevent overvotes. That is true, but modern PCOS can also alert a voter to overvotes and undervotes, allowing a voter to eject the ballot and correct it. Other solutions, such as non-tabulating scanners that simply warn voters of overvotes and undervotes on hand-marked ballots, would be less risky than BMDs.
- **Bad ballot design.** Ill-designed paper ballots, just like ill-designed touchscreen interfaces, may lead to unintentional undervotes [19]. For instance, the 2006 Sarasota, Florida, touchscreen ballot was badly designed. The 2018 Broward County, Florida, opscan ballot was badly designed: it violated three separate guidelines from the EAC's 2007 publication, "Effective Designs for the Administration of Federal Elections, Section 3: Optical scan ballots." [27] In both of these cases (touchscreens in 2006, hand-marked optical-scan in 2018), undervote rates were high. The solution is to follow standard, published ballot-design guidelines and other best practices, both for touchscreens and for hand-marked ballots [3, 19].
- **Low-tech paper-ballot fraud.** All paper ballots, however they are marked, are vulnerable to *loss*, *ballot-box stuffing*, *alteration*, and *substitution* between the

[20] "During the recount, the Coleman and Franken campaigns initially challenged a total of 6,655 ballot-interpretation decisions made by the human recounters. The State Canvassing Board asked the campaigns to voluntarily withdraw all but their most serious challenges, and in the end approximately 1,325 challenges remained. That is, approximately 5 ballots in 10,000 were ambiguous enough that one side or the other felt like arguing about it. The State Canvassing Board, in the end, classified all but 248 of these ballots as votes for one candidate or another. That is, approximately 1 ballot in 10,000 was ambiguous enough that the bipartisan recount board could not determine an intent to vote." [1] See also [20]

[21] We have found that some local election officials consider marks to be ambiguous if *machines* cannot read the marks. That is a different issue from *humans* not being able to interpret the marks. Errors in machine interpretation of voter intent can be dealt with by manual audits: if the reported outcome is wrong because machines misinterpreted handmade marks, a RLA has a known, large chance of correcting the outcome.

[22] Better designed user interfaces (UI) might reduce the error rate for machine-marked ballots below the historical rate for DREs; however, UI improvements cannot keep BMDs from printing something other than what the voter is shown on the screen.

time they are cast and the time they are recounted. That's why it is so important to make sure that ballot boxes are always in multiple-person (preferably bipartisan) custody at any time when they are handled, and that appropriate physical security measures are in place. Strong, verifiable chain-of-custody protections are essential.

Hand-marked paper ballots are vulnerable to alteration by anyone with a pen. Both hand-marked and BMD-marked paper ballots are vulnerable to substitution: anyone who has poorly supervised access to a legitimate BMD during election day can create fraudulent ballots, not necessarily to deposit them in the ballot box immediately (in case the ballot box is well supervised on election day) but with the hope of substituting it later in the chain of custody.[23]

All those attacks (on hand-marked and on BMD-marked paper ballots) are fairly low-tech. There are also higher-tech ways of producing ballots indistinguishable from BMD-marked ballots for substitution into the ballot box, where there is inadequate chain-of-custody protection.

- **Accessible voting technology.** If everyone voted on a BMD, it would guarantee that an accessible device had been set up in the polling place for all voters who needed one. But this is not a good reason to adopt BMDs for *all* voters. Among other things, it would expose all voters to the security flaws described above, decreasing public confidence in the entire election. Some accessibility advocates argue that requiring disabled voters to use BMDs compromises their privacy since hand marked ballots are easily distinguishable from machine marked ballots. This argument has been undercut by the availability in the marketplace of BMDs that mark ballots that cannot easily be distinguished from hand marked ballots. Other advocates object to the idea that disabled voters must use a different method of marking ballots, arguing that their rights are thereby violated. Both HAVA and ADA require accommodations for voters with physical and cognitive impairments, but neither law requires that those accommodations must be used by all voters. To best enable and facilitate participation by all voters, each voter should be provided with a means of casting a vote best suited to their abilities.
- **Ballot printing costs.** Preprinted optical-scan ballots cost 20–50 cents each.[24] Blank cards for BMDs cost up to 15 cents each, depending on the make and model of BMD.[25] But optical-scan ballots must be preprinted for as many vot-

[23] Some BMDs print a bar-code indicating when and where the ballot was produced, but that does not prevent such a substitution attack against currently EAC-certified, commercially available BMDs. We understand that systems under development might make ballot-substitution attacks against BMDs more difficult.

[24] Single-sheet (one- or two-side) ballots cost 20-28 cents, double-sheet ballots needed for elections with many contests, up to 50 cents.

[25] Ballot cards for ES&S ExpressVote cost about 15 cents. New Hampshire's (One4All / Prime III)

ers as *might* show up, whereas blank BMD cards are consumed in proportion to how many voters *do* show up. The Open Source Election Technology Institute (OSET) conducted an independent study of total life cycle costs[26] for hand-marked paper ballots and BMDs in conjunction with the 2019 Georgia legislative debate regarding BMDs [21]. OSET concluded that, even in the most optimistic (i.e., lowest cost) scenario for BMDs and the most pessimistic (i.e, highest cost) scenario for hand-marked paper ballots and ballot-on-demand (BOD) printers—which can print unmarked ballots as needed—the total lifecycle costs for BMDs would be higher than the corresponding costs for hand marked paper ballots.[27]

- **Vote centers.** To run a vote center that serves many election districts with different ballot styles, one must be able to provide each voter a ballot containing the contests that voter is eligible to vote in, possibly in a number of different languages. This is easy with BMDs, which can be programmed with all the appropriate ballot definitions. With preprinted optical-scan ballots, the PCOS can be programmed to *accept* many different ballot styles, but the vote center must still maintain *inventory* of many different ballots. BOD printers are another economical alternative for vote centers.[28]
- **Paper/storage.** BMDs that print summary cards rather than full-face ballots can save paper and storage space. However, many BMDs print full-face ballots, while many BMDs that print summary cards use thermal printers and paper that is fragile and can fade in a few months.[29]

Advocates of hand-marked paper ballot systems advance these additional arguments.

BMDs used by sight-impaired voters use plain paper that is less expensive.

[26]They include not only the cost of acquiring and implementing systems but also the ongoing licensing, logistics, and operating (purchasing paper stock, printing, and inventory management) costs.

[27]BOD printers currently on the market arguably are best suited for vote centers, but less expensive options suited for polling places could be developed. Indeed, BMDs that print full-face ballots could be re-purposed as BOD printers for polling place use, with modest changes to the programming.

[28]Ballot-on-demand printers *may* require maintenance such as replacement of toner cartridges. This is readily accomplished at a vote center with a professional staff. Ballot-on-demand printers may be a less attractive option for many small precincts on election day, where there is no professional staff—but on the other hand, they are less necessary, since far fewer ballot styles will be needed in any one precinct.

[29]The California Top-To-Bottom Review (TTBR) of voting systems found that thermal paper can also be covertly spoiled wholesale using common household chemicals https://votingsystems.cdn.sos.ca.gov/oversight/ttbr/red-diebold.pdf, last visited 8 April 2019. The fact that thermal paper printing can fade or deteriorate rapidly might mean it does not satisfy the federal requirement to preserve voting materials for 22 months. http://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title52-section20701&num=0&edition=prelim, last visited 8 April 2019.

- **Cost.** Using BMDs for all voters substantially increases the cost of acquiring, configuring, and maintaining the voting system. One PCOS can serve 1200 voters in a day, while one BMD can serve only about 260 [24]—though both these numbers vary greatly depending on the length of the ballot and the length of the day. OSET analyzed the relative costs of acquiring BMDs for Georgia's nearly seven million registered voters versus a system of hand marked paper ballots, scanners, and BOD printers [21]. A BMD solution for Georgia would cost taxpayers between 3 and 5 times the cost of a system based on hand marked paper ballots.
- **Mechanical reliability and capacity.** Pens are likely to have less downtime than BMDs. It is easy and inexpensive to get more pens and privacy screens when additional capacity is needed. If a precinct-count scanner goes down, people can still mark ballots with a pen; if the BMD goes down, voting stops. Thermal printers used in DREs with VVPAT are prone to jams; those in BMDs might have similar flaws.

These secondary pros and cons of BMDs do not outweigh the primary security and accuracy concern: BMDs, if hacked or erroneous, can change votes in a way that is not correctable. BMD voting systems are not contestable, defensible, or strongly software independent. Therefore, ballots cast by BMD cannot effectively be audited.

Barcodes

A controversial feature of some BMDs allows them to print 1-dimensional or 2-dimensional barcodes on the paper ballots. A 1-dimensional barcode resembles the pattern of vertical lines used to identify products by their universal product codes. A 2-dimensional barcode or QR code is a rectangular area covered in coded image *modules* that encode more complex patterns and information. BMDs print barcodes on the same paper ballot that contains human-readable ballot choices. Voters using BMDs are expected to verify the human-readable printing on the paper ballot card, but the presence of barcodes with human-readable text poses some significant problems.

- **Barcodes are not human readable.** The whole purpose of a paper ballot is to be able to recount (or audit) the *voters'* votes in a way independent of any (possibly hacked or buggy) computers. If the official vote on the ballot card is the barcode, then it is impossible for the voters to verify that the official vote they cast is the vote they expressed. Therefore, before a state even *considers* using BMDs that print barcodes (and we do not recommend doing so), the State must ensure by statute that recounts and audits are based *only* on the human-readable portion of

the paper ballot. Even so, audits based on untrusted paper trails suffer from the verifiability the problems we outlined above.

- **Ballot cards with barcodes contain two different votes.** Suppose a state does ensure by statute that recounts and audits are based on the human-readable portion of the paper ballot. Now a BMD-marked ballot card with both barcodes and human-readable text contains two different votes in each contest: the barcode (used for electronic tabulation), and the human-readable selection printout (official for audits and recounts). In few (if any) states has there even been a discussion of the legal issues raised when the official markings to be counted differ between the original count and a recount.
- **Barcodes pose technical risks.** Any coded input into a computer system—including wired network packets, WiFi, USB thumbdrives, *and barcodes*—pose the risk that the input-processing software can be vulnerable to attack via deliberately ill-formed input. Over the past two decades, many such vulnerabilities have been documented on *each* of these channels (including barcode readers) that, in the worst case, give the attacker complete control of a system.[30] If an attacker were able to compromise a BMD, the barcodes are an attack vector for the attacker to take over an optical scanner (PCOS or CCOS), too. Since it is good practice to close down all such unneeded attack vectors into PCOS or CCOS voting machines (e.g., don't connect your PCOS to the Internet!), it is also good practice to avoid unnecessary attack channels such as barcodes.

End-to-End Verifiable BMDs

In all BMD systems currently on the market, and in all BMD systems certified by the EAC, the printed ballot or ballot summary is the only channel by which voters can verify the correct recording of their ballots, independently of the computers. The analysis in this paper applies to all of those BMD systems.

There is a class of voting systems called "end-to-end verifiable" (E2E-V), which provide an alternate mechanism for voters to verify their votes [2]. Some E2E-V systems incorporate BMDs, for instance STAR-Vote[31] [5]. If such a voting system could

[30] An example of a barcode attack is based on the fact that many commercial barcode-scanner components (which system integrators use to build cash registers or voting machines) treat the barcode scanner using the same operating-system interface as if it were a keyboard device; and then some operating systems allow "keyboard escapes" or "keyboard function keys" to perform unexpected operations.

[31] The STAR-Vote system is actually a DRE+VVPAT system with a smart ballot box, rather than a BMD system: voters interact with a device that captures their votes electronically and prints a paper record that voters can inspect, but the electronic votes are held "in limbo" until the paper ballot is de-

be demonstrated to be contestable, defensible, and adequately usable by voters, then the analysis in this paper might not be applicable to such BMDs. No E2E-V systems are currently certified by the EAC, nor to our knowledge is any such system under review for certification, nor are any of the 5 major voting-machine vendors offering such a system for sale.[32]

Design Flaws in All-in-One BMDs

Some voting machines incorporate a BMD interface, printer, and optical scanner into the same cabinet. Other DRE+VVPAT voting machines incorporate ballot-marking, tabulation, and paper-printout retention, but without scanning.

These are often called "all-in-one" voting machines. Any such machine that includes *ballot marking* and *deposit into the ballot box* in the same paper path, is unsafe.

Using an all-in-one machine, the voter makes choices on a touchscreen or through a different accessible interface. When the selections are complete, the BMD prints the completed ballot for the voter to review and verify, before depositing the ballot in a ballot box attached to the machine.

- The ES&S ExpressVote (in all-in-one mode) allows the voter to mark a ballot by touchscreen or audio interface, then prints a paper ballot card and ejects it from a slot. The voter has the opportunity to review the ballot, then the voter redeposits the ballot into the same slot, where it is scanned and deposited into a ballot box.
- The ES&S ExpressVoteXL allows the voter to mark a ballot by touchscreen or audio interface, then prints a paper ballot (cash-register tape format) and displays it under glass. The voter has the opportunity to review the ballot, then the voter touches the screen to indicate "OK," and the machine pulls paper ballot up (still under glass) and into the integrated ballot box.
- The Dominion ImageCast Evolution (ICE) allows the voter to deposit a hand-marked paper ballot, which it scans and drops into the attached ballot box. *Or*, a voter can use a touchscreen or audio interface to direct the marking of a paper ballot, which the voting machine ejects through a slot for review; then the voter

posited in the smart ballot box. The ballot box does not read the votes from the ballot; rather, depositing the ballot tells the system that it has permission to cast the vote that it had already recorded from the touchscreen.

[32] Some vendors, notably Scytl, have sold systems advertised as E2E-V in other countries. Those systems were not in fact E2E-V. Moreover, serious security flaws have been found in their implementations. See, e.g., [16].

redeposits the ballot into the slot, where it is scanned and dropped into the ballot box.

In all three of these machines, the ballot-marking printer is in the same paper path as the mechanism to deposit marked ballots into an attached ballot box. This opens up a very serious security vulnerability: the voting machine can mark the paper ballot (to add votes or spoil already-cast votes) after the last time the voter sees the paper, and then deposit that marked ballot into the ballot box without the possibility of detection.

Vote-stealing software could easily be constructed that looks for *undervotes* on the ballot, and marks those unvoted spaces for the candidate of the hacker's choice. This is very straightforward to do on optical-scan bubble ballots (as on the Dominion ICE) where undervotes are indicated by no mark at all. On machines such as the ExpressVote and ExpressVoteXL, the normal software indicates an undervote with the words NO SELECTION MADE on the ballot summary card. Hacked software could simply leave a blank space there (most voters wouldn't notice the difference), and then fill in that space and add a matching bar code after the voter has clicked "cast this ballot."

An even worse feature of the ES&S ExpressVote and the Dominion ICE is the *auto-cast* configuration setting (in the manufacturer's standard software) that allows the voter to indicate, "don't eject the ballot for my review, just print it and cast it without me looking at it." If fraudulent software were installed in the ExpressVote, it could change *all* the votes of any voter who selected this option, because the voting machine software would know *in advance of printing* that the voter had waived the opportunity to inspect the printed ballot. We call this auto-cast feature "permission to cheat" [4].

Regarding these all-in-one machines, we conclude:

- Any machine with ballot printing in the same paper path with ballot deposit is not *software independent*; it is *not* the case that "an error or fault in the voting system software or hardware cannot cause an undetectable change in election results." Therefore such all-in-one machines do not comply with the VVSG 2.0 (the Election Assistance Commission's Voluntary Voting Systems Guidelines).
- All-in-one machines on which all voters use the BMD interface to mark their ballots (such as the ExpressVote and ExpressVoteXL) *also* suffer from the same serious problem as ordinary BMDs: most voters do not review their ballots effectively, and elections on these machines are not contestable or defensible.
- The auto-cast option for a voter to allow the paper ballot to be cast without human inspection is particularly dangerous, and states must insist that vendors disable or eliminate this mode from the software. However, even disabling the auto-cast feature does not eliminate the risk of undetected vote manipulation.

Remark. The Dominion ImageCast Precinct ICP320 is a precinct-count optical scanner (PCOS) that also contains an audio+buttons ballot-marking interface for disabled voters. This machine can be configured to cast electronic-only ballots from the BMD interface, or an external printer can be attached to print paper optical-scan ballots from the BMD interface. When the external printer is used, that printer's paper path is *not* connected to the scanner+ballot-box paper path (a person must take the ballot from the printer and deposit it into the scanner slot). Therefore this machine is as safe to use as any PCOS with a separate external BMD.

Conclusion

Ballot-Marking Devices produce ballots that do not necessarily record the vote expressed by the voter when they enter their selections on the touchscreen: hacking, bugs, and configuration errors can cause the BMDs to print votes that differ from what the voter entered and verified electronically. Furthermore, in cases where the BMD-marked paper ballot does not record the expressed vote, the election system overall is not *contestable* or *defensible*, meaning that errors in elections conducted on compromised BMDs cannot be reliably detected or corrected, and that election officials cannot provide convincing evidence that correct reported outcomes of elections conducted using BMDs are indeed correct. Therefore BMDs should not be used by voters who can use hand-marked paper ballots.

All-in-one voting machines, that combine ballot-marking and ballot-box-deposit into the same paper path, are even worse. They have all the disadvantages of BMDs (they are neither contestable or defensible), and they can mark the ballot after the voter has inspected it. Therefore they are not even *software independent*, and should not be used by those voters who are capable of marking, handling, and visually inspecting a paper ballot.

When computers are used to record votes, the original transaction (the voter's expression of the votes) is not documented in a verifiable way.[33] When pen-and-paper is used to record the vote, the original expression of the vote *is* documented in a verifiable way (provided that secure chain of custody of paper ballots is maintained). Therefore, audits of elections conducted with BMDs cannot ensure that reported outcomes are correct, while audits of elections conducted with hand-marked paper ballots, counted by optical scanners, can.

[33] It is conceivable that cryptographic protocols used in E2E-V systems could be used to create BMD-based systems that are contestable and defensible, but no such system exists, nor, to our knowledge, has such a design been worked out in principle.

References

[1] A.W. Appel. Optical-scan voting extremely accurate in Minnesota. *Freedom to Tinker*, January 2009. https://freedom-to-tinker.com/2009/01/21/optical-scan-voting-extremely-accurate-minnesota/.

[2] A.W. Appel. End-to-end verifiable elections. *Freedom to Tinker*, November 2018. https://freedom-to-tinker.com/2018/11/05/end-to-end-verifiable-elections/.

[3] A.W. Appel. Florida is the Florida of ballot-design mistakes. *Freedom to Tinker*, November 2018. https://freedom-to-tinker.com/2018/11/14/florida-is-the-florida-of-ballot-design-mistakes/.

[4] A.W. Appel. Serious design flaw in ESS ExpressVote touchscreen: "permission to cheat". *Freedom to Tinker*, September 2018. https://freedom-to-tinker.com/2018/09/14/serious-design-flaw-in-ess-expressvote-touchscreen-permission-to-cheat/.

[5] J. Benaloh, M. Byrne, B. Eakin, P. Kortum, N. McBurnett, O. Pereira, P.B. Stark, , and D.S. Wallach. Star-vote: A secure, transparent, auditable, and reliable voting system. *JETS: USENIX Journal of Election Technology and Systems*, 1:18–37, 2013.

[6] R. K. Bothwell, K.A. Deffenbacher, and J.C. Brigham. Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, 72:691–695, 1987.

[7] Election Assistance Commission. Voluntary voting systems guidelines 2.0, September 2017. https://www.eac.gov/assets/1/6/TGDC_Recommended_VVSG2.0_P_Gs.pdf.

[8] K. Deffenbacher. Eyewitness accuracy and confidence: Can we infer anything about their relation? *Law and Human Behavior*, 4:243–260, 1980.

[9] R. DeMillo, R. Kadel, and M. Marks. What voters are asked to verify affects ballot verification: A quantitative analysis of voters' memories of their ballots, November 2018. https://ssrn.com/abstract=3292208.

[10] S.L. Desmarais, T.L. Nicholls, J. D. Read, and J. Brink. Confidence and accuracy in assessments of short-term risks presented by forensic psychiatric patients. *The Journal of Forensic Psychiatry & Psychology*, 21(1):1–22, 2010.

[11] D. Dunning, D.W. Griffin, J.D. Milojkovic, and L. Ross. The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, 58:568–581, 1990.

[12] S.P. Everett. *The Usability of Electronic Voting Machines and How Votes Can Be Changed Without Detection*. PhD thesis, Rice University, 2007.

[13] A.J. Feldman, J.A. Halderman, and E.W. Felten. Security analysis of the Diebold AccuVote-TS voting machine. In *2007 USENIX/ACCURATE Electronic Voting Technology Workshop (EVT 2007)*, August 2007.

[14] Verified Voting Foundation. The verifier – polling place equipment – november 2018, November 2018. https://www.verifiedvoting.org/verifier/.

[15] D. Kahnemann. *Thinking, fast and slow*. Farrar, Straus and Giroux, 2011.

[16] S. J. Lewis, O. Pereira, and V. Teague. Ceci n'est pas une preuve: The use of trapdoor commitments in Bayer-Groth proofs and the implications for the verifiabilty of the Scytl-SwissPost Internet voting system, 2019. https://people.eng.unimelb.edu.au/vjteague/UniversalVerifiabilitySwissPost.pdf.

[17] M. Lindeman and P.B. Stark. A gentle introduction to risk-limiting audits. *IEEE Security and Privacy*, 10:42–49, 2012.

[18] National Academies of Sciences, Engineering, and Medicine. *Securing the Vote: Protecting American Democracy*. The National Academies Press, Washington, DC, September 2018.

[19] L. Norden, M. Chen, D. Kimball, and W. Quesenbery. Better Ballots, 2008. Brennan Center for Justice, http://www.brennancenter.org/publication/better-ballots.

[20] Office of the Minnesota Secretary of State. Minnesota's historic 2008 election, 2009. https://www.sos.state.mn.us/media/3078/minnesotas-historic-2008-election.pdf.

[21] E. Perez. Georgia state election technology acquisition: A reality check. OSET Institute Briefing, March 2019. https://trustthevote.org/wp-content/uploads/2019/03/06Mar19-OSETBriefing_GeorgiaSystemsCostAnalysis.pdf.

[22] K. Rayner and M.S. Castelhano. Eye movements during reading, scene perception, and visual search, 2009. Q J Experimental Psychology, 2009, August 62(8), 1457-1506.

[23] R.L. Rivest and J.P. Wack. On the notion of software independence in voting systems, July 2006. http://vote.nist.gov/SI-in-voting.pdf.

[24] Election Systems and Software. State of Georgia Electronic Request for Information New Voting System Event Number: 47800-SOS0000035, 2018. http://sos.ga.gov/admin/files/ESS%20RFI%20-%20Final%20-%20Redacted.pdf.

[25] P.B. Stark. Conservative statistical post-election audits. *Annals of Applied Statistics*, 2:550–581, 2008.

[26] P.B. Stark. Risk-limiting post-election audits: P-values from common probability inequalities. *IEEE Transactions on Information Forensics and Security*, 4:1005–1014, 2009.

[27] U. S. Election Assistance Commission. Effective designs for the administration of federal elections, June 2007. https://www.eac.gov/assets/1/1/EAC_Effective_Election_Design.pdf.

[28] J.T. Wixted and G.L. Wells. The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, 2017.